WHAT IS CLAIMED:

1. A method for reducing memory latency in a multi-node architecture, comprising:

receiving a speculative memory read request at a home node before results of a cache coherence protocol are determined; and

initiating a read to memory to complete the speculative memory read request.

2. The method of claim 2, further comprising:

buffering results of the read to memory.

3. The method of claim 2, further comprising:

dropping the results of the read to memory on a buffer full condition or if a cancel command is received.

4. The method of claim 3, further comprising:

if a confirm command is received after results of the read to memory are dropped, initiating a second read to memory to complete a memory read request.

5. The method of claim 4, further comprising:

forwarding results of the second read to memory to a requester.

6. The method of claim 3, further comprising:

if a confirm command is received before results of the speculative read are dropped, forwarding the results of the read to memory to a requester.

7. The method of claim 6, wherein the speculative memory read request is issued by the requesting node.

8. The method of claim 6, further comprising:

receiving the results of the read to memory at the coherence agent; and

forwarding the results of the read to memory to the requesting node.

9. A method for reducing memory latency, comprising:

issuing a speculative memory read request to a home node before results of a cache coherence protocol are determined; and

initiating the cache coherence protocol.

10. The method of claim 9, further comprising:

updating a memory status relating to the results in a table after the results of the cache coherence protocol are determined.

11. The method of claim 9, wherein initiating the cache coherence protocol comprising:

initiating a status look-up to determine the caching status of the requested memory.

12. The method of claim 11, further comprising:

issuing a confirm command to the home node if the caching status is determined to be in an invalid state or shared state.

13. The method of claim 11, further comprising:

snooping a node with the exclusive copy of the requested memory cached.

14. The method of claim 13, further comprising :

determining whether the exclusive copy of the requested memory is clean or dirty.

15. The method of claim 14, further comprising:

issuing a confirm command to the home node if the exclusive copy of the requested memory is clean.

16. The method of claim 14, further comprising:

issuing a cancel command to the home node if the exclusive copy of the requested memory is dirty.

17. The method of claim 13, further comprising:

receiving a snoop result, wherein the snoop result includes a copy of the requested memory; and

updating a memory status relating to the requested memory in a table.

18. The method of claim 17, further comprising:

receiving the requested memory; and

forwarding the requested memory to a requesting node.

19. A home node for responding to read requests in a multi-node architecture including a plurality of nodes, the home node comprising:

a processor;

a memory; and

a node controller coupled to the processor and memory, the node controller adapted to:

receive a speculative memory read request from a requester in the multi-node architecture before a cache coherence protocol is resolved, and

initiate a read to memory to complete the speculative memory read request.

20. The home node of claim 19 further comprising:

a buffer adapted to buffer the results of the read to memory.

21. The home node of claim 20, wherein the results of the read from memory are dropped from the buffer on a buffer full condition or upon receiving a cancel command.

22. The home node of claim 20, wherein the node controller responsive to a confirm is adapted to forward the results of the read to memory to the requester.

23. The home node of claim 20, wherein the node controller responsive to a cancel command is adapted to drop the data specified by the speculative read command.

24. A system comprising:

a node including a node controller adapted to control a plurality of processors resident in the node, wherein the node controller adapted to receive a speculative read request before results of a coherence protocol are determined and the node controller adapted to read data specified by the speculative read command from memory; and

a coherence agent coupled to the at least one node, the coherence agent including a coherence controller adapted to determine the results of the coherence protocol and adapted to forward a cancel command or a confirm command to the node after the results of the coherence protocol are determined.

25. The system of claim 24, wherein the node controller responsive to the confirm command issued by the coherence controller is adapted to send the data read from memory to the coherence controller.

26. The system of claim 24, wherein the node controller responsive to the cancel command issued by the coherence controller is adapted to drop the data read from memory.

27. The system of claim 24, further comprising:

a requesting node adapted to send a data read request to request data identified by a memory address included in the data read request.

28. The system of claim 27, wherein the speculative read request is sent by the requesting node.

29. The system of claim 24, wherein the speculative read request is sent by the switching agent.